

## PATENT COOPERATION TREATY

From the INTERNATIONAL BUREAU

PCT

## NOTIFICATION OF ELECTION

(PCT Rule 61.2)

To:

Commissioner  
US Department of Commerce  
United States Patent and Trademark  
Office, PCT  
2011 South Clark Place Room  
CP2/5C24  
Arlington, VA 22202  
ETATS-UNIS D'AMERIQUE  
in its capacity as elected Office

Date of mailing (day/month/year) 21 March 2001 (21.03.01)	
International application No. PCT/US00/14250	Applicant's or agent's file reference 10003628-1
International filing date (day/month/year) 24 May 2000 (24.05.00)	Priority date (day/month/year) 24 May 1999 (24.05.99)
Applicant KRAUSE, Michael et al	

1. The designated Office is hereby notified of its election made:

☒ in the demand filed with the International Preliminary Examining Authority on:  
22 December 2000 (22.12.00)

☐ in a notice effecting later election filed with the International Bureau on:

2. The election ☒ was  
☐ was not

made before the expiration of 19 months from the priority date or, where Rule 32 applies, within the time limit under Rule 32.2(b).

The International Bureau of WIPO 34, chemin des Colombettes 1211 Geneva 20, Switzerland Facsimile No.: (41-22) 740.14.35	Authorized officer Juan Cruz Telephone No.: (41-22) 338.83.38
---	---

# PATENT COOPERATION TREATY

## PCT

### INTERNATIONAL PRELIMINARY EXAMINATION REPORT

(PCT Article 36 and Rule 70)

10  
RECD 24 DEC 2001

09/980761

Applicant's or agent's file reference 10003628-1	FOR FURTHER ACTION See Notification of Transmittal of International Preliminary Examination Report (Form PCT/IPEA/416)	
International application No. PCT/US00/14250	International filing date (day/month/year) 24 MAY 2000	Priority date (day/month/year) 24 MAY 1999
International Patent Classification (IPC) or national classification and IPC IPC(7): HO2H 3/05; G06F 15/16 and US Cl.: 714/4; 709/201		
Applicant HEWLETT-PACKARD COMPANY		

1. This international preliminary examination report has been prepared by this International Preliminary Examining Authority and is transmitted to the applicant according to Article 36.

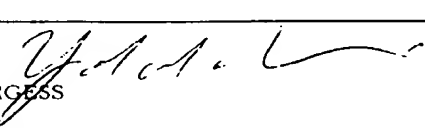
2. This REPORT consists of a total of 3 sheets.

☒ This report is also accompanied by ANNEXES, i.e., sheets of the description, claims and/or drawings which have been amended and are the basis for this report and/or sheets containing rectifications made before this Authority. (see Rule 70.16 and Section 607 of the Administrative Instructions under the PCT).

These annexes consist of a total of 4 sheets.

3. This report contains indications relating to the following items:

- I ☒ Basis of the report
- II ☐ Priority
- III ☐ Non-establishment of report with regard to novelty, inventive step or industrial applicability
- IV ☐ Lack of unity of invention
- V ☒ Reasoned statement under Article 35(2) with regard to novelty, inventive step or industrial applicability; citations and explanations supporting such statement
- VI ☐ Certain documents cited
- VII ☐ Certain defects in the international application
- VIII ☐ Certain observations on the international application

Date of submission of the demand  22 DECEMBER 2000	Date of completion of this report  31 JULY 2001
Name and mailing address of the IPEA/US Commissioner of Patents and Trademarks Box PCT Washington, D.C. 20231	Authorized officer  GLENTON BURGESS
Facsimile No. (703) 305-3230	Telephone No. (703) 305-3900

## INTERNATIONAL PRELIMINARY EXAMINATION REPORT

International application No.

PCT/US00/14250

**I. Basis of the report****1. With regard to the elements of the international application:\***

- ☐ the international application as originally filed
- ☒ the description:  
pages 1-43 , as originally filed  
pages NONE , filed with the demand  
pages NONE , filed with the letter of \_\_\_\_\_
- ☒ the claims:  
pages NONE , as originally filed  
pages NONE , as amended (together with any statement) under Article 19  
pages NONE , filed with the demand  
pages 44-47 , filed with the letter of 18 JUNE 2001
- ☒ the drawings:  
pages 1-13 , as originally filed  
pages NONE , filed with the demand  
pages NONE , filed with the letter of \_\_\_\_\_
- ☒ the sequence listing part of the description:  
pages NONE , as originally filed  
pages NONE , filed with the demand  
pages NONE , filed with the letter of \_\_\_\_\_

**2. With regard to the language, all the elements marked above were available or furnished to this Authority in the language in which the international application was filed, unless otherwise indicated under this item.**

These elements were available or furnished to this Authority in the following language \_\_\_\_\_ which is:

- ☐ the language of a translation furnished for the purposes of international search (under Rule 23.1(b)).
- ☐ the language of publication of the international application (under Rule 48.3(b)).
- ☐ the language of the translation furnished for the purposes of international preliminary examination (under Rules 55.2 and/or 55.3).

**3. With regard to any nucleotide and/or amino acid sequence disclosed in the international application, the international preliminary examination was carried out on the basis of the sequence listing:**

- ☐ contained in the international application in printed form.
- ☐ filed together with the international application in computer readable form.
- ☐ furnished subsequently to this Authority in written form.
- ☐ furnished subsequently to this Authority in computer readable form.
- ☐ The statement that the subsequently furnished written sequence listing does not go beyond the disclosure in the international application as filed has been furnished.
- ☐ The statement that the information recorded in computer readable form is identical to the written sequence listing has been furnished.

**4. ☒ The amendments have resulted in the cancellation of:**

- ☒ the description, pages NONE
- ☒ the claims, Nos. NONE
- ☒ the drawings, sheets/~~fig~~ NONE

**5. ☐ This report has been drawn as if (some of) the amendments had not been made, since they have been considered to go beyond the disclosure as filed, as indicated in the Supplemental Box (Rule 70.2(c)).\*\***

\* Replacement sheets which have been furnished to the receiving Office in response to an invitation under Article 14 are referred to in this report as "originally filed" and are not annexed to this report since they do not contain amendments (Rules 70.16 and 70.17).

\*\*Any replacement sheet containing such amendments must be referred to under item 1 and annexed to this report.

# INTERNATIONAL PRELIMINARY EXAMINATION REPORT

International application No.

PCT/US00/14250

## **V. Reasoned statement under Article 35(2) with regard to novelty, inventive step or industrial applicability; citations and explanations supporting such statement**

### **1. statement**

Novelty (N)	Claims <u>1-23</u>	YES
	Claims <u>NONE</u>	NO
Inventive Step (IS)	Claims <u>1-23</u>	YES
	Claims <u>NONE</u>	NO
Industrial Applicability (IA)	Claims <u>1-23</u>	YES
	Claims <u>NONE</u>	NO

### **2. citations and explanations (Rule 70.7)**

Claims 1-23 meet the criteria set out in PCT Article 33(2)-(4), because the prior art does not teach or fairly suggest multi-unicast operations enabled with the send work queue at the source endnode and destination endnodes and end-to-end contexts at the source endnode and destination endnodes participating in a multicast group.

----- NEW CITATIONS -----  
NONE

WHAT IS CLAIMED IS:

1. A distributed computer system comprising:
  - a source endnode participating in a multicast group and including:
    - a source process which produces message data;
    - a send work queue having work queue elements that describe the message data for multicasting;
  - multiple destination endnodes participating in a multicast group, each destination endnode including:
    - a destination process;
    - a receive work queue having work queue elements that describe where to place incoming message data;
  - communication fabric providing communication between the source endnode and the multiple destination endnodes; and
  - multiple end-to-end contexts, each end-to-end context having a portion storing state information at the source node and a portion storing state information at a corresponding one of the destination endnodes to ensure the reception and sequencing of message data multicast from the source endnode to the corresponding one of the destination endnodes, wherein a reliable multicast comprises a series of replicated unicasts of message data through the send work queue and each of the end-to-end contexts portions at the source endnode to the receive work queue and the corresponding end-to-end context portion at each of the destination endnodes.
2. The distributed computer system of claim 1 wherein the source endnode includes a network interface controller which packetizes the message data into frames.
3. The distributed computer system of claim 2 wherein the destination endnodes each include a network interface controller which acknowledges receipt of frames multicast from the source endnode.
4. The distributed computer system of claim 3 wherein the network interface controller and the end-to-end context portion in each destination

endnode ensure strong ordering of received frames multicast from the source endnode, such that the frames are received in a same defined order as transmitted from the source endnode.

5. The distributed computer system of claim 3 wherein the source endnode retransmits frames that are not successively acknowledged in the reliable multicast.
6. The distributed computer system of claim 2 wherein the network interface controller in the source endnode includes hardware which replicates frames to be provided in the series of unicasts.
7. The distributed computer system of claim 2 wherein the source endnode includes software verbs which perform the series of unicasts as a series of individual sequenced message send operations.
8. The distributed computer system of claim 1 wherein changes in composition of the endnodes participating in the multicast group are communicated to all endnodes participating in the multicast group.
9. The distributed computer system of claim 1 wherein the source endnode and each destination endnode maintains a list of destination addresses for all other endnodes participating in the multicast group.
10. The distributed computer system of claim 3 wherein the network interface controller in each destination endnode generates cumulative acknowledgments.
11. The distributed computer system of claim 3 wherein the network interface controller in each destination endnode generates acknowledgments on a per frame basis.
12. The distributed computer system of claim 3 the network interface controller of the source endnode includes a completion processing unit which

gathers acknowledgements from the destination endnodes and completes frame operation by informing the source process of an operation status of multicast frames.

13. The distributed computer system of claim 12 wherein the source endnode further comprises:

a completion queue containing information related to completed work queue elements, wherein the completion processing unit communicates with the source process via the completion queue.

14. The distributed computer system of claim 12 wherein the completion processing unit informs the source process which destination processes, if any, did not receive multicast frames.

15. The distributed computer system of claim 12 wherein the completion processing unit includes an acknowledgement counter which counts acknowledgements received from the corresponding destination endnodes in the multicast group indicating that the corresponding destination endnode has received a frame multicast from the source endnode.

16. The distributed computer system of claim 15 wherein completion processing unit generates a completion event to the source process when the acknowledgement counter indicates that a predetermined percentage of the destination endnodes in the multicast group have acknowledged the multicast frame has been received.

17. The distributed computer system of claim 15 wherein completion processing unit generates a completion event to the source process when the acknowledgement counter indicates that all of the destination endnodes in the multicast group have acknowledged the multicast frame has been received.

18. The distributed computer system of claim 12 wherein the completion processing unit includes a bit-mask array which assigns a unique bit for each destination endnode in the multicast group and clears each bit as a corresponding

acknowledgment is received from the corresponding destination endnode in the multicast group indicating that the corresponding destination endnode has received a frame multicast from the source endnode.

19. The distributed computer system of claim 18 wherein the completion processing unit generates a completion event to the source process when the bit-mask array has a predetermined percentage of bits cleared in the bit-mask array indicating that that a predetermined percentage of the destination endnodes in the multicast group have acknowledged the multicast frame has been received.

20. The distributed computer system of claim 18 wherein the completion processing unit generates a completion event to the source process when the bit-mask array has all bits cleared in the bit-mask array indicating that that all of the destination endnodes in the multicast group have acknowledged the multicast frame has been received.

21. The distributed computer system of claim 12 wherein the completion processing unit includes a timing window, wherein expiring of the timing window without necessary conditions for a completion event for a corresponding multicast frame occurring indicates that any missing acknowledgments are to be tracked and resolved.

22. The distributed computer system of claim 1 wherein a given process joins the multicast group by performing a multicast join operation.

23. The distributed computer system of claim 1 wherein a given process leaves the multicast group by performing a multicast leave operation.





patent (AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, GW, ML, MR, NE, SN, TD, TG).

*For two-letter codes and other abbreviations, refer to the "Guidance Notes on Codes and Abbreviations" appearing at the beginning of each regular issue of the PCT Gazette.*

**Published:**

— *With international search report.*